

National Center for Education Statistics Disclaimer

The information and opinions published here are the product of the International Indicators of Education Systems project's Network A and do not necessarily represent the policy or views of the Department of Education or the National Center for Education Statistics.

NETWORK A MEETING RECORD

Network A Plenary Meeting October 14-15, 2004, El Escorial, Madrid, Spain

Participants

Helmut Bachmann, Austria
Christiane Blondin, Belgium (French)
Luc Van de Poele, Belgium (Flemish)
Jerry Mussio, Canada
Pavla Zieleniecova, Czech Republic
Jørn Skovsgaard, Denmark
Jorma Kuusela, Finland
Thierry Rocher, France
Zsuzsa Hamori-Vaczy, Hungary
Julius Bjornsson, Iceland
Ryo Watanabe, Japan
Iris Blanke, Luxembourg
Paul van Oijen, Netherlands
Lynne Whitney, New Zealand
Marit Granheim, Norway
Glória Ramalho, Portugal
Vladislav Rosa, Slovak Republic
Guillermo Gil, Spain
Anita Wester, Sweden
Erich Ramseier, Switzerland
Müfide Çaliskan, Turkey
Jason Tarsh, United Kingdom
Elois Scott, United States
Eugene Owen, Network A Chair
Jay Moskowitz, Network A Secretariat
Maria Stephens, Network A Secretariat
Andreas Schleicher, OECD Secretariat

Observers

Jochen Schweitzer, Germany
Flora Gil Traver, Spain
Ramón Pajares Box, Spain

Regrets

Wendy Whitham, Australia
Jürgen Horschinegg, Austria
Botho Priebe, Germany
Evangelia Varnava-Skoura, Greece
Gerry Shiel, Ireland
Chiara Croce, Italy
Myungjoon Lee, Korea
Felipe Martinez Rizo, Mexico
Fernando Cordova, Mexico

Summary of Major Outcomes

Network A had four main points of business at the meeting, discussing: activities related to the Task Force on Teaching and Learning, indicators for *Education at a Glance* (EAG) 2005, a proposal for a technical review of PISA, and planning for subgroup activities. The major outcomes were:

- With regard to the task force on teaching and learning, the Network reiterated its concerns about a teacher survey linked to PISA and requested that these comments be represented at the PGB meeting. The Network A Secretariat will keep members informed about developments with the proposal and next steps with regard to a longer-term strategy.

- With regard to indicators for EAG 2005, members will review the indicators and send any comments in writing to the Network A Secretariat. The Secretariat will revise the indicators in accordance with those comments, as well as reintroduce an indicator utilizing TIMSS eighth-grade data, make more prominent the trend information from the TIMSS fourth-grade results, and consider the indicator on students' beliefs about themselves as learners as the lowest priority.
- With regard to a technical review of PISA, the Network A Secretariat will take nominations from members as well as seek additional nominations in order to ensure representation of a range of technical expertise, and will be back in touch with the Network with a proposed panel and next steps to proceed.
- With regard to subgroups, chairs will proceed to work with group members on the scopes of work discussed at the meeting. The Network A Chair will be in touch with the subgroup chairs to ensure that there is no overlap among the three groups and coordination is achieved where possible.
- The Network A Secretariat will circulate the documents requested at the meeting, including: the results from the poll regarding strategic development of PISA and information on developments with the adult competencies study. Luc van de Poele will let members know how to log on to the on-line forum set up for the subgroup on development.
- The next meeting will be in Stockholm on March 2-4, 2005. Subgroup meetings will be held on the afternoon of the 2nd and the plenary session will take place on the 3rd and 4th. The Network meeting will be followed by the PISA PGB meeting on March 7-9.

Welcome and Introduction

Eugene Owen opened the Network A meeting and introduced Alejandro Tiana Ferrer to make an official welcome on behalf of the Spanish authorities. He then announced new participants, who would be joining the meeting later, including Júlíus Björnsson from Iceland, Marit Granheim from Norway, and Müfide Çaliskan from Turkey, as well as participants from the Spanish INES team, Flora Gil and Ramón Pajares. He also gave regrets for: Australia, Greece, Ireland, Italy, Korea, and Mexico. He then went over the agenda, noting the reversal of the teaching discussion and the INES updates, and it was adopted. [The minutes of the prior meeting in Lucerne were approved without changes in a later session.]

Discussion on Task Force on Teaching and Learning

Eugene opened the meeting with an update and discussion on the activity related to developing a strategy for collecting information on teachers, teaching and learning. He introduced the papers in the briefing book, which included a cover paper from the OECD Secretariat and two expert reviews that had been commissioned on the what information could or could not be gained from a survey of teachers in PISA schools on their attitudes, working conditions, and the instructional environment.

In responding to a question from Luc Van de Poele about what the scope of the discussion should and could be, Eugene noted that it was important for the Network to go on record again with their position on the proposal for a teacher survey in their role in providing advice and

guidance to the PGB on issues related to development of PISA. The importance of countries speaking with one voice, with the multiple groups (e.g., Networks, PBG, Education Committee, CERI Governing Board, INES SMG) commenting on the proposal, was mentioned by Eugene and other members.

Members then had a lengthy discussion, commenting on the advisability of a teacher survey in PISA schools and the content of the expert papers. Pavla Zieleniecova noted that in some countries, linking activities such as a teacher survey to PISA increased the possibility for securing funding for them. Jerry Mussio suggested that part of a response could include an analysis of the extent to which the proposed survey meets or does not meet the original expectations for development work in the area of teachers and teaching. Luc noted that while both expert papers concluded that a teacher survey could not provide insight on teacher effectiveness, they diverged on whether or not it could provide insight on school effectiveness. In doing so, he noted that the cover paper did not treat this divergence objectively, rather seeming to favor one viewpoint over the other. Ryo Watanabe provided some background on the INES SMG's recent recommendation regarding the activity: he said that the SMG had advised proceeding with development work because the area is one of such high priority and had recommended a cross-ISCED level survey with a sub-sample in PISA schools.

Jørn Skovsgaard suggested that one strategy to address the concern that information from a teacher survey in PISA schools would be used to make statements about effectiveness would be to report results separately from PISA assessment results. Elois Scott was concerned that even such efforts at the international level would not prevent the eventual linking of the information on teachers and PISA student outcomes and moreover thought that, in the United States, the inclusion of additional components in PISA could have a detrimental effect on school participation rates, which are already the biggest challenge in national-level implementation. Anita Wester said that the expert reviews were helpful and that, from the Swedish perspective, they validated their reluctance to support a teacher survey in PISA schools. She also noted the potential negative impacts on response rates in her country and noted that in many countries, the students in PISA schools would have a very short history with the teachers in the survey, decreasing the value of information to be gained.

Christiane Blondin noted that it seems like there's been a shift to "what we can do in the short-term" instead of a more considered long-term approach. She also expressed concerns over the possible misuse of results as well as the ability to develop and implement a well designed survey even of a limited scope in time for PISA 2006 or shortly thereafter. Jay Moskowitz also expressed concern that the long-term perspective was being lost.

Paul van Oijen noted that the reporting on the proposal lacked information from the experiences of other large-scale assessment programs with teacher surveys, for example: the perhaps mixed positive and negative impacts on participation rates from the inclusion of teacher surveys. He noted interest in a study that would utilize other methodologies such as classroom observation. Jason Tarsh also noted the lack of drawing upon the TIMSS experience in the reviews, proposal and related papers, though noting some sympathy for the need for information from a survey of teachers. Elois noted that a teacher survey was somewhat more of a fit in IEA studies because of the assessments' basis in school curricula.

Lynne Whitney commented that although there was a great interest in information on teachers and teaching, it would be preferable to take time to design a study that would seek the information of greatest interest (i.e., related to teaching practice). She thought that if a teacher survey were conducted in PISA schools it would be difficult to counter expectations that it could provide information on school or teacher effectiveness. Marit Granheim also emphasized the special role of Network A in looking longer term and noted that the nature of PISA would not well support a teacher survey. Jochen Schweitzer noted the country-specific element of the debate, giving the example that in Germany, the interest a teacher survey stems from the nature of the system and the important policy question of how teachers address students' needs or if they are tracked to other parts of the system.

Luc pointed out Klieme's recommendation that mathematics teachers as the most appropriate subjects and the poor fit of a survey to PISA, which in 2006 will focus on scientific literacy. He also noted Klieme's recommendation to seek information on both deep and surface teaching, which might not be possible under the timeline. Jerry noted that, although there is an obvious gap in indicators on information related to teaching, an important first task would also be to examine the relationship among existing process and outcome indicators.

Marit also shared her experiences in recent work at the World Bank, suggesting that there are many areas of student learning (such as learning to learn) to explore before tackling complex issues related to linking teaching and learning outcomes. Erich Ramseier shared concerns that it would not be feasible to develop a survey in time for PISA 2006 and, in addition, the emphasis the proposal puts on triangulation among teacher, principal, and student surveys would require additional changes to the latter two surveys, which also may not be feasible under the timeline. He expressed a preference for taking a longer-term perspective and exploring longitudinal methods.

Guillermo Gil stated that that he was not in favor of a teacher questionnaire in PISA or linking the results with the current PISA design and also preferred a more comprehensive study of teaching separate from PISA. Flora Gil added that a teacher survey would not be able to fulfill what the policy makers would expect from it – namely, explanations for student performance – and that it would be better to take profit from current observational studies and the German teacher study in reconsidering the long-term perspective. Christiane suggested that the only way to avoid the results of a teacher survey being linked to PISA student results would be to use a national sample at a particular or cross-ISCED level.

Jason urged caution with the argument that PISA's nature as a yield study would not support a teacher survey *per se*, arguing that the concept of yield and the distinction of PISA from other curriculum-based large-scale assessment has not been well explained to date. Iris Blanke said that in Luxembourg, they were able to get some information on school environment by administering the principal survey to teachers and suggested this as a possible alternative to a teacher survey as proposed. Jorma Kuusela shared that, in the Finnish experience, in which there have been many surveys and evaluations of teachers in conjunction with student evaluations that had not yielded valuable results. He suggested that the importance of a teacher survey might also relate to the amount of variation between schools in different countries; in the Finnish context,

there are small differences between schools and thus arguably not as many differences in teaching and instructional environments. Following on this comment, Paul called attention to research on school effectiveness that shows small effect sizes of teacher/teaching factors compared to students' backgrounds and earlier educational experiences.

At this point in the conversation, Eugene Owen offered a draft statement summarizing the Network's position and suggested that the Network return to the discussion on the following day.

[This discussion continued on the second day of the meeting.] The original draft statement read as follows: "Network A recognizes the desire to obtain more information about teachers, teaching, and learning. We consider the proposed teacher survey as a limited first step. Network members appreciate the contributions of Jürgen Baumert and Eckhard Klieme in clarifying the complexity and limits that a teacher survey will provide. The majority of Network A members prefers a separate survey of teachers, independent of PISA and they question the ability to launch a survey connected to PISA by 2006. Network A strongly supports working towards an internationally comparative student of teacher characteristics and teaching processes, which is not possible through a questionnaire, nor linked to PISA."

Most members were supportive of the sentiments in the statement, though a few were not. They requested that the statement provide a more detailed overview of the reasons for the argument against a teacher survey, as well as note the dissention from the three members that were more open to the possibility. The Network A Secretariat will keep members informed about developments with the proposal and next steps with regard to a longer-term strategy.

Discussion on EAG Indicators

Maria Stephens then took the floor to introduce the draft indicators for EAG 2005. She reminded members of the many open questions following the last meeting in Switzerland, relating both to the possibility of expanding the number of indicators included in EAG and which indicators should be included. Over the summer, the Network A Secretariat clarified with the OECD Secretariat the number of indicators that could be included—the usual 5-6—and developed two alternative proposals for indicators. Both proposals included a core set of indicators that consisted of the key results from PISA, with one proposal focusing on additional results from PISA, mainly in mathematical literacy and the other proposal focusing more on composite indicators (e.g., gender) and trend indicators. She noted that neither proposal included information from TIMSS eighth-grade, because of concerns about explaining the differences between TIMSS and PISA.

Maria noted that members' responses slightly favored the first package of indicators, with the focus on PISA mathematics results and so the draft indicators follow along that outline. The first indicator utilizes IEA data from TIMSS 2003 at the fourth-grade level, focusing on the 11 OECD countries for which there are data. The remaining indicators draw from PISA, and Maria noted that since they were modified from the draft PISA international report, which is still evolving, the indicators were as well and would be checked and finalized in accordance with the final report. The second indicator, which is the most different from the international report—as it focuses more on overall versus subscale results—is on mathematical literacy. The third indicator presents the reading and scientific literacy results. Trend information from these two indicators is

kept minimal since there are only two data points and is presented in sidebars. The fourth indicator presents the results from the problem solving assessment. The fifth indicator attempts to answer countries' interests in including an indicator dealing with students' attitudes and engagement and focuses on the results from three scales on students' beliefs about themselves as learners, where the results were among the most interesting. The final indicator presents information on between- and within-school variation. In conclusion, she asked members both to prioritize among the indicators and to share any suggestions for their finalization.

Jason shared the first comments, who: noted his concern at the lack of TIMSS eighth-grade data; suggested that the trend information in the TIMSS fourth-grade data is arguably the most interesting and should be more prevalent; and expressed a preference for an indicator on the SES gradients over one on students' beliefs about themselves as learners. Jerry supported Jason's first comment, arguing that the Network should take the opportunity to signal their belief in the importance of having a diversity of assessments and to be more reflective of the complexity of the task of measuring outcomes and the need for different sources of information. Marit further supported this idea, saying that the community should take every opportunity to present information to help audiences understand the differences between the different available measures.

Luc also suggested that, since the PISA indicators now include information from non-OECD countries, the TIMSS indicators might do the same. He also noted that the method of significance testing for the TIMSS indicator should be clarified (and be done consistently with PISA methods) and the multiple comparisons double-checked in light of this. Christiane noted that the trend information should be either presented in the main text or removed (rather than displayed in the text box, preferring these devices for providing clarifying information, such as on the differences between PISA and TIMSS mathematical content). [In a subsequent session,] it also was suggested that the explanatory information on the mathematics subscales be expanded. On the larger question of prioritization among indicators, several members suggested that the indicator of least interest was the one related to students' beliefs about themselves as learners.

Because of time limitations in this session [and the following day], it was suggested that any further comments on the TIMSS indicator and the other PISA indicators should be sent to the Network A Secretariat. Summarizing, Jay said the Network A Secretariat would revise the indicators to expand the information on TIMSS fourth-grade trends, reintroduce an indicator utilizing TIMSS eighth-grade data, consider the indicator on beliefs as the lowest priority, and take other forthcoming comments into account.

Updates on INES and PISA

Andreas Schleicher then took the floor to provide an update on activities in INES and PISA. [The full presentation is attached for information.]

In response to the presentation, Luc requested that information on the continuing development of the adult competencies activity be disseminated to Network A and noted somewhat of an incongruence between the open issues still to be decided (e.g., related to populations, scope) and the topics on the agenda for the next meeting, suggesting that additional discussion time may be needed for development and consensus building. Andreas noted that information on the adult

competencies study is currently disseminated to National Coordinators but that it could be shared with Network A members as well. Several members also reiterated the importance of improving coordination between OECD and IEA, and Andreas responded that it is up to countries to first determine what their information needs are collectively and then explore the appropriate sources to fill those gaps. It also was suggested that OECD recirculate the poll asking countries about issues for the development of PISA after 2009. In particular, it was suggested that the issue of cycles could be revisited with more careful analysis of the pros and cons of the various options. Members also requested that results from the original poll be circulated to Network A.

Reports from Subgroups

Each of the three subgroup chairs then reported to members on the outcomes of their planning meetings. The presentations and/or summary notes from each of these are attached. In response to a question about avoiding overlap among the groups (e.g., in polling members or establishing on-line forums), Eugene said that he would work with the subgroup chairs to coordinate across groups.

Discussion on Technical Review of PISA

Eugene then introduced a proposal to undertake a technical review of PISA. He noted that the proposal focused on reviewing methodological issues related to establishing stable trend lines. He distinguished this from a longer-term, broader evaluation of PISA, which still needs to be thought out and discussed with the appropriate subgroup.

Erich supported the importance of undertaking a technical review and suggested that it should focus on contrasting and analyzing different approaches, since PISA has only been exposed to one contractor. He suggested that, more than geographical diversity, reviewers should represent a diversity of methodological expertise and experience. Jay suggested that the reviewers might be asked to undertake simulations of PISA 2006 data under different models and approaches, in addition to providing theoretical critiques. Paul van Oijen thought it would be useful to first identify what issues PISA was currently confronting and specifically why such a technical review was necessary. In terms of the reviewers' products, it also was suggested that in addition to the technical response, there should be some interpretation accessible to lay audiences.

Iris asked about the process of nomination, and Eugene suggested that countries are invited to send nominations to the Network A Secretariat. Jerry suggested that other methods of nominations, such as through research by the Secretariat would also be appropriate, to ensure a range of expertise, as had been emphasized by Erich. On the issue of diversity, there was a long discussion on how wide-ranging the opinions should be—i.e., should strong critics of international assessments be included? Some members suggested that it could be useful bringing a perspective about other methods of data analysis particularly at the item level, and others thought that it would be harmful because of—at least in the example being discussed—the staunch position against the usefulness of IRT scaling in general. In the end, it was decided that critical literature could be presented as background information for the experts charged with undertaking the review.

Updates on Other Projects

In the final session, Eugene Owen gave a brief overview of the current status of IEA studies. He noted:

- With regard to PIRLS: The last meeting of the PIRLS 2006 National Research Coordinators (NRCs) took place in Bratislava, Slovakia in March 2004. The purpose of that meeting was to finalize the reading passages for the assessment, and begin to develop assessment items related to the passages. The next meeting will be in Miami, Florida in November, where the focus will be on reviewing instruments for the field test next spring. About 50 countries have signed up to participate in PIRLS 2006, the framework for which was just publicly released.
- With regard to TIMSS: The last meeting was in June 2004 in Santiago, Chile for countries to go over the weighted, scaled results. Results will be released on December 14, 2004 for both mathematics and science and both fourth-grade and eighth-grade. The technical report also will be released at the time, though the user guide will follow in March 2005. The database and user guide will follow in March 2005. Released item sets were published in July 2004. Fifty countries participated in the eighth-grade assessment and 27 participated in the fourth-grade assessment. The first meeting of the TIMSS 2007 NRCs will be February 6-12, 2005 in Egypt, in conjunction with database training.

Conclusion and Next Meeting

[Earlier in the meeting,] Jochen Schweitzer said goodbye to the Network, announcing that Botho Priebe would be the new German representative and would attend the next plenary meeting. He thanked the Network for the fruitful and enjoyable collaborations over the years, and the Network thanked him in return for his contributions.

The next meeting of Network A will be in Stockholm on March 2-4, 2005, in advance of the PISA PGB meeting on March 7-9, 2005. Finally, Eugene thanked members for their cooperation and hard work, and the meeting was adjourned.